

24/2/2013

Μοντέλα Πολλαπλής Παλινδρόμησης

Πέτρος Ρούσσο

Πρόγραμμα Ψυχολογίας, ΦΠΨ, ΕΚΠΑ

ΕΙΣΑΓΩΓΙΚΑ

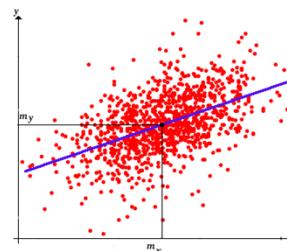
Ορολογία

- Προβλεπτικές μεταβλητές ή παράγοντες (predictors)
- Μεταβλητή κριτήριο (criterion)
- [Απλή](#) και πολλαπλή παλινδρόμηση
- Ευθύγραμμη παλινδρόμηση
- Απλή πολλαπλή παλινδρόμηση
- Βηματική πολλαπλή παλινδρόμηση
- Ιεραρχική πολλαπλή παλινδρόμηση

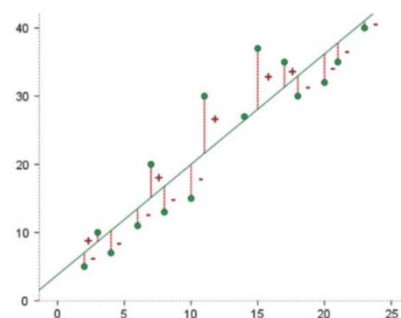
Απλή παλινδρόμηση

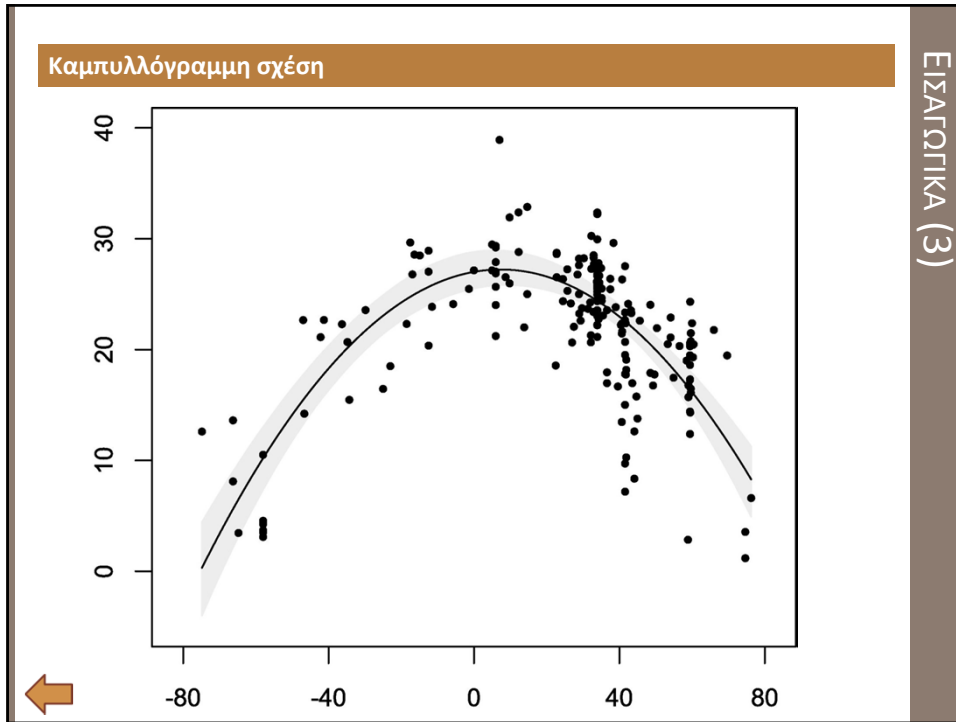
- Διάγραμμα σκεδασμού (scatterplot)
- Ευθύγραμμη ή [καμπυλόγραμμη](#) σχέση;
- Η γραμμή παλινδρόμησης
- Εξίσωση της ευθείας γραμμής:
 $Y = \alpha + bX$
- Συντελεστές παλινδρόμησης
- α = σταθερός όρος (constant)
- b = κλίση (slope)
- Τιμές πρόβλεψης (predicted values)
- Υπόλοιπα (residuals)
- Συντελεστής προσδιορισμού (r^2)

Η γραμμή παλινδρόμησης



Υπόλοιπα (Residuals)



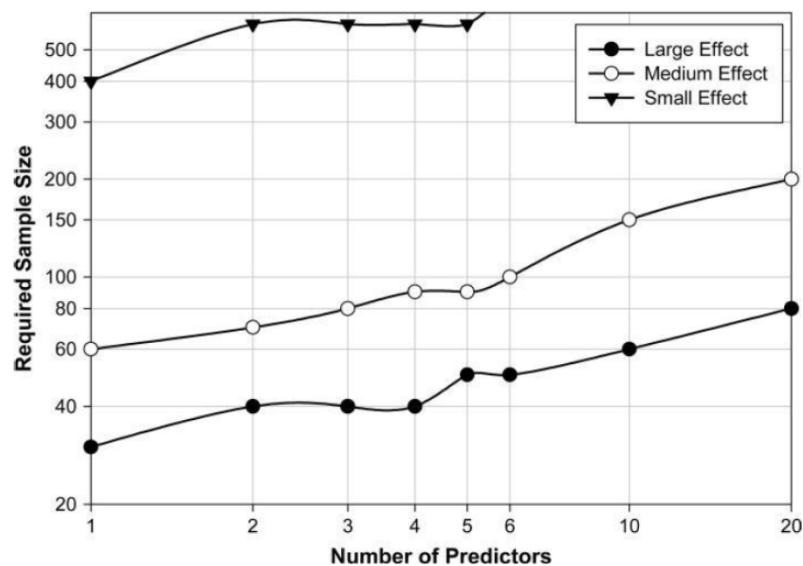


ΠΡΟΫΠΟΘΕΣΕΙΣ ΕΦΑΡΜΟΓΗΣ ΚΡΙΤΗΡΙΟΥ

Μέγεθος του δείγματος

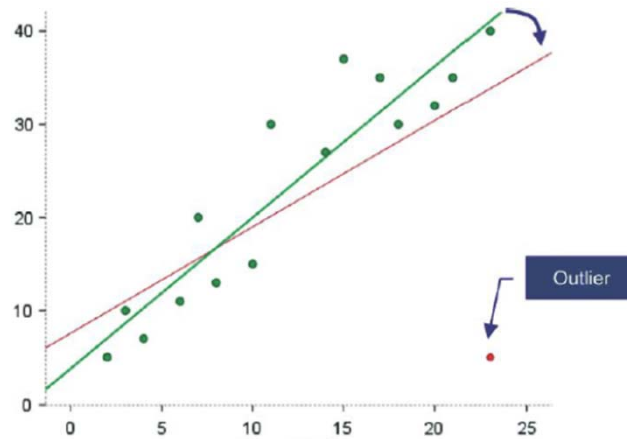
- Στην περίπτωση της απλής πολλαπλής παλινδρόμησης – τουλάχιστον 20 περιπτώσεις ανά προβλεπτικό παράγοντα (ελάχιστο 10)
- Harris (1985) – το N θα πρέπει να ξεπερνά τους προβλεπτικούς παράγοντες κατά 50
- Cohen & Cohen (1975) – Power: μία προβλεπτική μεταβλητή με ισχύ 0,80 & συντελεστή συσχέτισης 0,30 → N = 124
- Βηματική πολλαπλή παλινδρόμηση – διπλάσιες περιπτώσεις ανά προβλεπτική μεταβλητή
- Όταν υπάρχουν πάρα πολλές περιπτώσεις είναι ευκολότερο να πάρουμε στατιστικώς σημαντικά αποτελέσματα...
- Ορθότερο να κάνουμε ανάλυση ισχύος!

Μέγεθος του δείγματος



Έκτοπες ή ακραίες τιμές

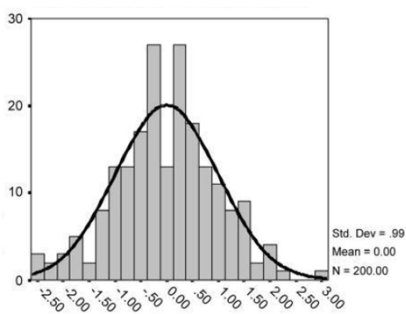
- Έλεγχος για την ύπαρξη **έκτοπων τιμών**, οι οποίες έχουν μεγάλη επίδραση στο αποτέλεσμα της ανάλυσης...



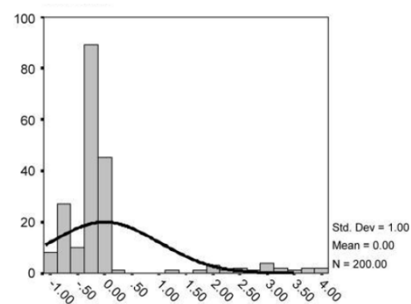
ΠΡΟΫΠΟΘΕΣΕΙΣ ΕΦΑΡΜΟΓΗΣ ΚΡΙΤΗΡΙΟΥ (3)

Κανονικότητα των σφαλμάτων

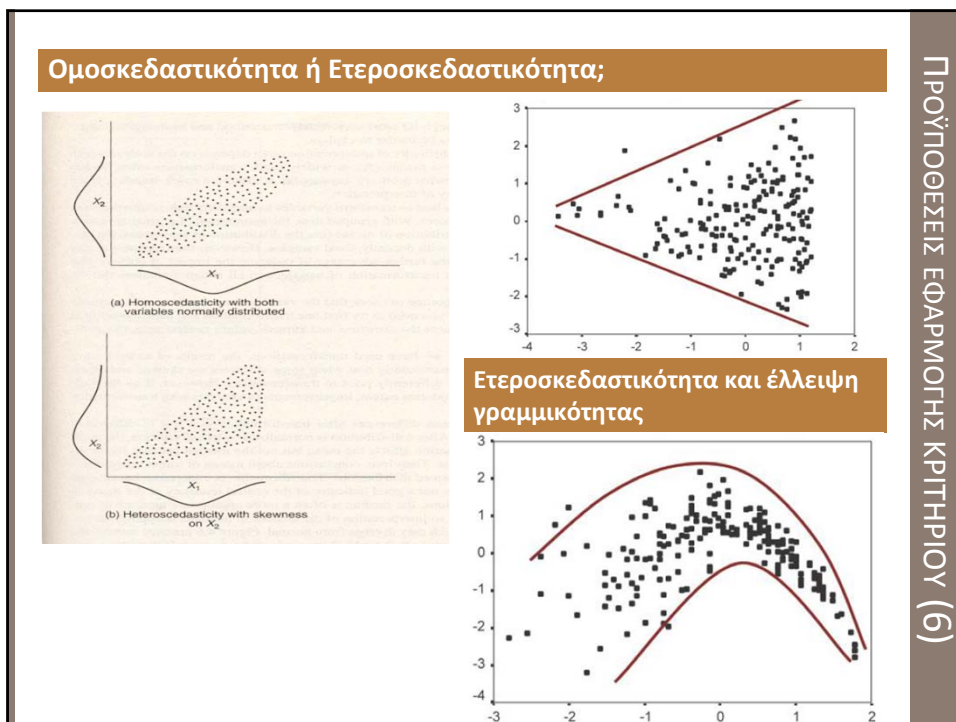
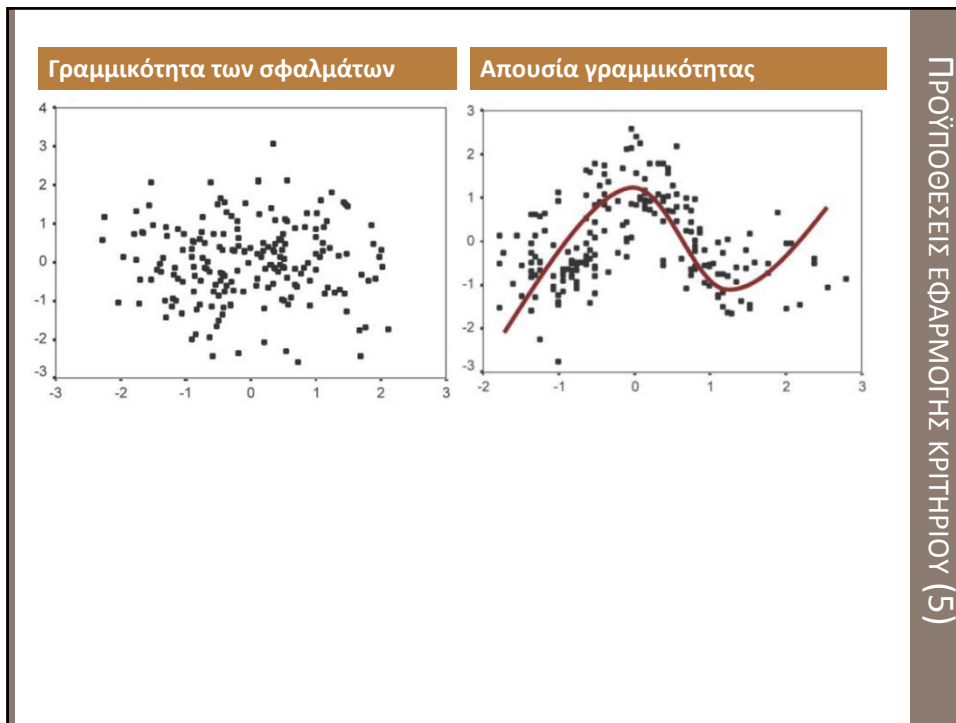
- Κανονικότητα των σφαλμάτων (random error)
 - Μελέτη των υπολοίπων μέσω του ιστογράμματος των τυποποιημένων υπολοίπων:



Μη κανονικότητα?



ΠΡΟΫΠΟΘΕΣΕΙΣ ΕΦΑΡΜΟΓΗΣ ΚΡΙΤΗΡΙΟΥ (4)



Άλλες προϋποθέσεις

- Η κλίμακα μέτρησης της μεταβλητής κριτήριο πρέπει να είναι ίσων διαστημάτων
- **Υπαρξη συσχέτισης** μεταξύ της μεταβλητής κριτήριο και των προβλεπτικών μεταβλητών
- Ανεξαρτησία των σφαλμάτων: για δύο οποιεσδήποτε παρατηρήσεις, τα υπόλοιπα πρέπει να είναι ανεξάρτητα
 - Ελέγχεται με το κριτήριο Durbin-Watson (η τιμή του πρέπει να είναι γύρω στο 2)
- **Πολυσυγγραμμικότητα** (multicollinearity) →
 - συντελεστές συσχέτισης των προβλεπτικών μεταβλητών $>0,80$?
 - Αν ΝΑΙ, αφαίρεση της μίας μεταβλητής...
 - Διαγνωστικοί δείκτες (collinearity statistics)
 - Δείκτης ανοχής (Tolerance)
 - Παράγοντας πληθωριστικής διασποράς (VIF)

Πίνακας 1.

Συντελεστές συσχέτισης μεταξύ της μεταβλητής κριτήριο και των προβλεπτικών μεταβλητών.

Μετρήσεις	1	2	3	4	5
1. Ανάγνωση	-	.75***	.46***	.42***	-.36**
2. Αντιμεταθέσεις		-	.61***	.31**	-.54***
3. Ομοιοκαταληξίες			-	.34**	-.39***
4. Raven				-	-.22*
5. RAN					-

***p < .001; **p < .01; * p < .05



Δυο λόγια για τα ερευνητικά δεδομένα του workshop**▪ Δείγμα:**

N = 86 παιδιά (n = 42 αγόρια & n = 44 κορίτσια)

Ηλικία σε μήνες – $M = 81.73$, $SD = 3.39$.

▪ Μετρήσεις:

- Ανάγνωση – αριθμός ορθά αναγνωσμένων λέξεων (κριτήριο / DV)
- Τεστ Raven (μη λεκτική / αφηρημένη συλλογιστική ικανότητα)
- RAD (αριθμός γνωστών αντικειμένων που ονομάζονται ορθά)
- Μετρήσεις φωνημικής ενημερότητας:
 - Ομοιοκαταληξίες
 - Αντιμεταθέσεις – RED → BOX (τα παιδιά ακούν δύο λέξεις και πρέπει να δημιουργήσουν δύο νέες λέξεις αντιμεταθέτοντας τα αρχικά φωνήματα: π.χ., BED → ROX)

Η ΔΙΑΔΙΚΑΣΙΑ ΣΤΟ SPSS

Οι εντολές

- Από το αρχικό μενού επιλέξτε [Analyze → Regression → Linear]
- Μετακινήστε την εξαρτημένη μεταβλητή της έρευνάς σας στο πλαίσιο [Dependent] και τις ανεξάρτητες μεταβλητές στο πλαίσιο [Independent(s)]
- Πρέπει να επιλέξουμε μέθοδο εισαγωγής των προβλεπτικών μεταβλητών στο μοντέλο παλινδρόμησης: Επιλογή [Method:]
 - Ως αρχική επιλογή έχει τη μέθοδο [Enter]. Εάν πατήσετε πάνω στο κουμπί με την ένδειξη [Enter] θα αναδυθεί ένα νέο μενού, στο οποίο περιλαμβάνονται τέσσερις ακόμη μέθοδοι (Stepwise, Remove, Backward, και Forward)
- Επιλογή [Statistics...] και στο πλαίσιο διαλόγου [Linear Regression: Statistics]
 - Στις δύο προεπιλογές [Estimates] και [Model Fit] προσθέστε τις επιλογές: [R square change], [Descriptives] και [Collinearity diagnostics].
- Επιλογή [Plots...] και στο πλαίσιο διαλόγου [Linear Regression: Plots]
 - Για το γράφημα με την κατανομή των υπολοίπων, επιλέξτε [Histogram] από την ενότητα [Standardized Residual Plots], και τοποθετήστε τη μεταβλητή *ZRESID στο πλαίσιο με την ένδειξη Y: και τη μεταβλητή *ZPRED στο πλαίσιο με την ένδειξη X:

Πίνακας 2.

Ιεραρχική πολλαπλή παλινδρόμηση με την ανάγνωση ως κριτήριο.

Προβλεπτική μεταβλητή	B	SE B	β
Μοντέλο 1			
Αντιμεταθέσεις	1.20	.15	.75***
Ομοιοκαταληξίες	.03	.29	.01
Μοντέλο 2			
Αντιμεταθέσεις	1.15	.14	.71***
Ομοιοκαταληξίες	-.13	.28	-.04
Raven	.55	.20	.21**
Μοντέλο 3			
Αντιμεταθέσεις	1.21	.16	.75***
Ομοιοκαταληξίες	-.12	.28	-.03
Raven	.57	.20	.22**
RAN	.04	.04	.08

$R^2 = .56$, Adjusted $R^2 = .55$ (Μοντέλο 1); $R^2 = .60$, Adjusted $R^2 = .59$ (Μοντέλο 2); $R^2 = .61$, Adjusted $R^2 = .59$ (Μοντέλο 3); R^2 change = .56 (Μοντέλο 1); R^2 change = -.04 (Μοντέλο 2); R^2 change = .00 (Μοντέλο 3)

*** $p < .001$; ** $p < .01$